GAUSS AND SAS FOR RECOVERING INTERBLOCK AND INTERVARIETY INFORMATION

by

Walter T. Federer and Russell D. Wolfinger
Biometrics Unit, Cornell U., and SAS Institute, Inc.

ABSTRACT

Augmented experiment designs are useful for screening large numbers of new treatments. There are two types of treatments involved, checks or standards and new treatments. The former are fixed effects while the new are usually considered to be random effects. Interblock (or interrow and intercolumn) information needs to be recovered for check treatments. Interreplicate, interblock (or interrow and intercolumn), and intervariety (new treatment) information needs to be recovered for the new treatments. It is possible to do this using available software packages. GAUSS and SAS PROC GLM and PROC MIXED procedures are presented for recovering these types of information.

BU-1384-M October 1996

In the Technical Report Series of the Biometrics Unit, Cornell University, Ithaca, New York 14853.

1. Introduction

The class of augmented experiment designs (Federer, 1956b, 1961, 1991; Federer *et al*, 1975, 1975) was introduced to replace the systematically placed check arrangement for screening new genotypes. Usually material is limited necessitating that one experimental unit be allocated to each new genotype, or if the material is not limited, it may not be desirable to include a treatment more than once. An appropriate statistical analysis was presented by Federer (1996). The analysis takes the random nature of the new genotypes and blocking effects into account. Augmented experiment designs have several advantages over the systematic check arrangement.

An augmented experiment design (AED) is a useful design for screening new treatments (genotypes, insecticides, herbicides, drugs, etc.), where n the number of new treatments is large, even in the hundreds and thousands. An AED is constructed by selecting a standard experiment design (ED) for the c check or standard treatments. The design could be a randomized complete block design (RCBD), an incomplete block design (BIBD, PBIBD), a row-column design (RCD), or a resolvable row-column (RRCD) such as a lattice square design. Then, the rb blocks (b incomplete blocks per complete block) are enlarged so that n/rb new treatments can be included (randomly) in each incomplete block. Each block will contain s = k + n/rb experimental units (eus), i.e., the standard incomplete block design is augmented with n/rb new treatments. Although the new treatments usually appear once in an experiment, there is nothing to preclude the experimenter from using as many experimental units as desired for a new treatment. This will only affect the analysis. The analyses described herein are for one replicate for each new treatment.

In the above it was assumed that n/rb, a constant, new treatments would be allocated in each block. This need not be the case as illustrated for the augmented BIBD (ABIBD) described below. There are c = 4 checks, k = 2 checks per incomplete block, s = 3 entries per incomplete block, b = 2 incomplete blocks per replicate, and r = 3 replicates for the standard design for checks. Suppose that the checks are numbered 12, 13, 14, and 15 and that there are n = 11 new treatments. A schematic design plan is:

Incomplete block	Replicate 1	Replicate 2	Replicate 3
1	1 2 12 13	3 4 12 14	5 6 12 15
2 = k	7 8 14 15	9 10 13 15	11 x 13 14

where x denotes a blank (or one of the new treatments could be replicated).

In the above it may be noted that the check treatments are given the largest numbers. Numbering in this manner results in a more informative analysis using the SAS software package. In PROC GLM, the highest numbered class variable is set equal to zero in obtaining the solutions for effects. Since it is not reasonable to set any of the new treatments equal to zero under a random effects model and since all the new could be worse or all could be better than the checks, numbering in this manner by-passes this. Under the model that the sum of the check effects equals zero, the solutions will be the treatment effect minus the highest numbered treatment effect and the standard errors listed in the SAS output will actually be standard errors of a difference of two effects (or means).

AEDs were introduced to replace the systematically spaced single check arrangement, where the check was included in every kth eu., e.g., every third, fourth, ..., tenth eu. This arrangement allocates a high proportion of eus to the check treatment and hence, is an inefficient procedure (Federer, 1956a). Yates (1936) showed that the number of check eus should be approximately $n^{1/2}$, where n is the number of eus allocated to new treatments. The systematic arrangement does not provide an estimate of the error variance for the new treatments and for comparing them with the check, whereas an AED does. More than one check is available for comparison with new treatments in an AED.

Since the new treatments should ordinarily be considered to be random effects and since every AED is an incomplete block design with respect to the new treatments, both interblock and intervariety information needs to be recovered in order to provide an efficient statistical analysis for the data. The calculations for performing such analyses may be easily obtained from available software packages. We demonstrate how to perform these calculations using the GAUSS and SAS PROC GLM and PROC MIXED packages. A small numerical example is used to illustrate the programs.

2. Statistical Analyses for Augmented Block Designs

Suppose an incomplete block design is selected as the standard design for the checks and the b incomplete blocks are enlarged to include n_i ,$\Sigma_{i=1 \text{ to } rb}\, n_i = n$, new treatments in incomplete block hi. Then, a response model equation may be

$$Y_{hij} = \mu + \rho_h + \beta_{hi} + \tau_j + \varepsilon_{hij}, \tag{1}$$

where μ is a general effect, ρ_h is the hth replicate effect distributed with mean zero and variance $\sigma_\rho{}^2$, β_{hi} is the ith incomplete block effect in complete block (replicate) h distributed with mean zero and variance $\sigma_\beta{}^2$, h = 1, ..., r, i = 1, ..., b, τ_j is the effect of the jth treatment and let γ_j is the effect for the jth check, j = n+1, ..., n+c, and η_j is the effect of new treatment j distributed with $\eta_.$ and variance $\sigma_\eta{}^2$, j = 1, ..., n, and ε_{hij} are random error effects distributed with unknown mean and variance $\sigma_\varepsilon{}^2$.

In order to simplify the presentation, we use the responses Y_{hij} minus the mean effect from check responses only, i.e., Y_{hij} - $\hat{\mu_c}$ values. When the new treatments occur only once in the experiment and when the incomplete block sizes are all equal, a matrix form of the above response equations is

$$\begin{bmatrix} bsI_r & RB_{r\times rb} & RC_{r\times c} & RN_{r\times n} \\ RB'_{rb\times r} & k\,I_{rb} & C_{rb\times c} & N_{rb\times n} \\ RC'_{c\times r} & C'_{c\times rb} & r\,I_c & 0_{c\times n} \\ RN'_{n\times r} & N'_{n\times rb} & 0_{n\times c} & I_n \end{bmatrix} \begin{bmatrix} \rho_{r\times 1} \\ \beta_{rb\times 1} \\ \gamma_{c\times 1} \\ \eta_{n\times 1} \end{bmatrix} = \begin{bmatrix} YR_{r\times 1} \\ YB_{rb\times 1} \\ YC_{c\times 1} \\ YN_{n\times 1} \end{bmatrix} \qquad (2)$$

where there are bs treatments in replicate h, I_x is an identity matrix of side x, **0** is a matrix of zeros, **RB** is a replicate by block incidence matrix, **RN** is a replicate by new treatment incidence matrix, **RC** is a replicate by check incidence matrix, **C** is the block by check incidence matrix, **N** is the block by new incidence matrix, ρ is a column vector of replicate effects, β is a column vector of incomplete block effects, γ is a column vector of check treatment effects, η is a column vector of new treatment effects, **YR** is a column vector of replicate totals minus the mean, **YB** is a vector of block totals adjusted for the mean, **YC** is a vector of check totals adjusted for the mean, and **YN** is the new treatment response minus mean effect. If the replicate sizes vary, then replace bs I_r with a diagonal matrix of replicate sizes on the diagonal. If the block sizes vary, simply replace k I_{rs} by a diagonal matrix with block sizes on the diagonal. If the new treatment occurs more than once, replace I_n with a diagonal matrix with the replicate numbers on the diagonal. Equation (2) is over-parameterized and replacing **RB** by **RB** - d **J** and **C** with (**C** - d **J**), e.g., where **J** is a matrix of ones and d is an appropriate scalar, results in unique solutions for the parameters of (2); this amounts to using the restriction that the sum of the incomplete block within replicate effects and the check effects sum to zero.

The intrablock (fixed effects) solutions for β, and γ, are

$$\hat{\beta} = \left[s\,I_{rb} - ((C - J)\ \ N)\begin{bmatrix} r\,I_c & 0 \\ 0' & I_n \end{bmatrix}^{-1}\begin{bmatrix} C' \\ N' \end{bmatrix}\right]^{-1}$$
$$\times\ [YB - ((C - J)\ \ N)]\begin{bmatrix} r\,I_c & 0 \\ 0' & I_n \end{bmatrix}^{-1}\begin{bmatrix} YC \\ YN \end{bmatrix} \qquad (3)$$

and

$$\hat{\gamma} = \left[r\mathbf{I}_c - ((\mathbf{C}' \;\; - \;\; \mathbf{J}') \;\; \mathbf{0}) \begin{bmatrix} s\,\mathbf{I}_{rb} & \mathbf{N} \\ \mathbf{N}' & \mathbf{I}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{C} \\ \mathbf{0}' \end{bmatrix} \right]^{-1}$$

$$\times \left[\mathbf{YC} \;\; - ((\mathbf{C} \;\; - \;\; \mathbf{J}') \;\; \mathbf{0})\right] \begin{bmatrix} s\,\mathbf{I}_{rb} & \mathbf{N} \\ \mathbf{N}' & \mathbf{I}_n \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{YB} \\ \mathbf{YN} \end{bmatrix} \qquad (4)$$

Since the new treatment effects are not orthogonal to either blocks or replicates, the form for $\hat{\eta}$ is more complex. It is

$$\hat{\eta} = \left[\mathbf{I}_n - (\mathbf{RN}' \;\; \mathbf{N}' \;\; \mathbf{0}') \begin{bmatrix} bs\,\mathbf{I}_r & (\mathbf{RB}\text{-}\mathbf{JR}) & (\mathbf{RC} - \mathbf{J}) \\ \mathbf{RB} & s\,\mathbf{I}_{rb} & (\mathbf{C} - \;\; \mathbf{J}) \\ \mathbf{RC}' & \mathbf{C}' & r\,\mathbf{I}_c \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{RN} \\ \mathbf{N} \\ \mathbf{0} \end{bmatrix} \right]^{-1}$$

$$\times \;\; [\mathbf{YN} \;\; - (\mathbf{RN}' \;\; \mathbf{N}' \;\; \mathbf{0}'] \begin{bmatrix} bs\,\mathbf{I}_r & (\mathbf{RB}\text{-}\mathbf{JR}) & (\mathbf{RC} - \mathbf{J}) \\ \mathbf{RB}' & s\,\mathbf{I}_{rb} & (\mathbf{C} - \;\; \mathbf{J}) \\ \mathbf{RC}' & \mathbf{C}' & r\,\mathbf{I}_c \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{YR} \\ \mathbf{YB} \\ \mathbf{YC} \end{bmatrix} \qquad (5)$$

where **JR** and **J** are the restriction matrices on solutions for blocks and checks, respectively, for estimated effects summing to zero..

To recover interblock information on the checks, replace k $\mathbf{I}_{rb}$ in (4) with $(k + \sigma_\varepsilon^2 / \sigma_\beta^2)\, \mathbf{I}_{rb}$, where estimates of the variance components σ_ε^2 and σ_β^2 are used. To recover interreplicate, interblock, and intervariety information on the new, replace bs $\mathbf{I}_r$ with $(bs + \sigma_\varepsilon^2 / \sigma_\rho^2)\mathbf{I}_r$, replace k $\mathbf{I}_{rb}$ with $(k + \sigma_\varepsilon^2 / \sigma_\beta^2)\, \mathbf{I}_{rb}$, and replace $\mathbf{I}_n$ with $(1 + \sigma_\varepsilon^2 / \sigma_\eta^2)\, \mathbf{I}_n$, where estimates of the variance components σ_ε^2, σ_ρ^2, σ_β^2 , and σ_η^2 are used. The GAUSS program as written, uses ANOVA solutions for the variance components whereas SAS PROC MIXED uses restricted maximum likelihood (REML) solutions.

The various sums of squares are computed as follows. The block (eliminating other effects), denoted as Type III, sum of squares is computed as $\hat{\beta}'$ times the second term in (3). The check (eliminating other effects) sum of squares is computed as $\hat{\gamma}'$ times the second term in (4). The new (eliminating other effects) sum of squares is computed as $\hat{\eta}'$ times the second term in (5).

The first terms in (3), (4), and (5) times σ_ε^2 are the variance-covariance matrices for the $\hat{\beta}$, $\hat{\gamma}$, and $\hat{\eta}$, respectively. When replacements are made as described above for recovering interreplicate, interblock and intervariety information, the variance-covariance matrices for the effects adjusted for these types of information result.

3. A Numerical Example

A simple example which exhibits the results of Section 2 is a BIBD with $c = 4$ checks, $r = 3$ replicates, and incomplete blocks of size $k = 2$. Then, to accommodate using $n = 6$ new treatments, each with one replicate, the incomplete block is enlarged to $s = 3$ eus. New treatments are numbered 1, 2, ..., $6 = n$ and check treatments are numbered 7, 8, 9, and $10 = n + c = v$. A numerical example (Table 1) of Y_{hij} values was constructed using the following values of the parameters:

$m = 10$	$\rho_1 = -5$	$\rho_2 = 0$	$\rho_3 = 5$	$\beta_{11} = 4$	$\beta_{12} = -4$
$\beta_{21} = 0$	$\beta_{22} = 0$	$\beta_{31} = -3$	$\beta_{32} = 3$	$\tau_1 = 10$	$\tau_2 = 0$
$\tau_3 = 0$	$\gamma_4 = -5$	$\tau_5 = 7$	$\tau_6 = -2$	$\tau_7 = -4$	$\tau_8 = 0$
$\tau_9 = 0$	$\tau_{10} = 4$	$\varepsilon_{117} = -1$	$\varepsilon_{118} = 1$	$\varepsilon_{129} = 1$	$\varepsilon_{1210} = -1$
$\varepsilon_{217} = 1$	$\varepsilon_{219} = -1$	$\varepsilon_{228} = -1$	$\varepsilon_{2210} = 1$		

The remaining ε_{hij}s are equal to zero. The fixed (intrablock) effects solutions obtained from the GAUSS and SAS programs must equal the above. The error sum of squares must equal 8 since there only the squares of 8 ± 1s. Constructing an example in this manner is a helpful device for evaluating the correctness of a program.

Table 1. Layout, responses, and totals for $c = 4$ checks in a BIBD with the incomplete blocks augmented with one new.

	Replicate 1				Replicate 2				Replicate 3			
				sum				sum				sum
	7	8	1		7	9	3		7	10	5	
	4	10	19	33	7	9	10	26	8	16	19	43
	9	10	2		8	10	4		8	9	6	
	2	4	1	7	9	15	5	29	18	18	16	52
Replicate total				40				55				95

Check totals: (7) 19 (8) 37 (9) 29 (10) 35
New totals: (1) 19 (2) 1 (3) 10 (4) 5 (5) 19 (6) 16

Note that checks are orthogonal to replicate effects but not to blocks. New are not orthogonal to either replicates or blocks. Owing to the orthogonality with replicates, the check sum of squares in Table 2 is that obtained for a RCBD on check yields. When the number of new per replicate is the same for all replicates, the contrast for check versus new is also orthogonal to replicate effects and is computed in the usual manner. This leaves only the among new (eliminating replicate effect) sum of squares to compute in order to partition the treatment (check and new) (eliminating replicate but ignoring block) sum of squares into its component parts as given in the ANOVA in the top part of Table 2. The calculations are carried to four decimals for the purposes of limiting rounding errors when comparing results with computer outputs. For this particular example, the sum of squares for treatments (ignoring replicate and block effects) is 363.7778, and the sum of squares for replicate (ignoring treatment) is 269.4444. The sum of squares for replicates from check responses only is 200. Thus, the correction for disproportionality is the difference between these two sums of squares or 69.4444. If the correction for disproportionality is subtracted from the among new (ignoring replicate and block) sum of squares, the sum of squares for among new (eliminating replicate and ignoring block) is obtained. Thus,

$$19^2 + 1^2 + 10^2 + 5^2 + 19^2 + 16^2 - (19 + 1 + 10 + 5 + 19 + 16)^2/6 - 69.4444$$
$$= 1104 - 4900/6 = 287.3333 - 69.4444 = 217.8889.$$

The sum of squares among new treatments (ignoring replicate and block effects) is 217.8889 + 69.4444 = 287. 3333. The contrast of check versus new is:

$$120^2/12 + 70^2/6 - (120 + 70)^2/18 = 11.1111.$$

The check sum of squares ignoring blocks is

$$(19^2 + 37^2 + 29^2 + 35^2)/3 - 120^2/12 = 65.3333.$$

These three sums of squares add to the treatment (ignoring block and replicate effects) sum of squares, i.e., 287.3333 + 11.1111 + 65.3333 = 363.7778. The sum of squares for treatment (eliminating replicate effect but ignoring block effect) is 363.7778 - 69.4444 = 217.8889 + 65.3333 + 11.1111 = 294.3333.

The ANOVA on check responses only is obtained in the usual manner for a BIBD. Block (eliminating check) sum of squares from check responses only is the same as when both check and new responses are used. Since there is only one response for each new treatment, this response can contribute nothing to obtaining solutions for the mean, replicate effects, and block effects. For this example on check responses only, there are $k = 2$ responses per incomplete block and for purposes of obtaining the expected value for blocks (eliminating all other effects) mean square, the incomplete block size for checks only is k = 2; the ANOVA expected value is $\sigma_\varepsilon^2 + k\,(r - 1)\,\sigma_\beta^2 / r$, given that the β_{hi} are random effects distributed with mean zero and estimated variance of $\sigma_\beta\hat{}^2 = 3(22.2222 - 2.6667)/4 = 14.6667$. The degrees of freedom for new treatment (eliminating all other effects) mean square is n - 1. The expected value of this mean square is $\sigma_\varepsilon^2 + \sigma_\eta^2$, given that the η_j are distributed with unknown mean $\eta_.$ and variance σ_η^2. A solution for σ_η^2 is $\sigma\hat{}_\eta^2 = 111.0794 / 5 - 2.6667 = 19.5492$. The estimated replicate component of variance $\sigma_\rho\hat{}^2$ is $(200 / 2 - 2(14.6667) - 2.6667) / 4 = 17.0000$. Using these solutions, we are now in a position to obtain new effects recovering interreplicate, interblock, and intervariety information as explained for equation (5). Also, interblock information is used to obtain the check effects as described for equation (4).

Following the analysis given in textbooks for a balanced incomplete block designed experiment, e.g., Federer (1955) and Cochran and Cox (19557), the treatment 7, 8, 9, and 10 means with recovery of interblock information are 6.06,10.28, 9.96, and 13.72, respectively. The corresponding effects are -3.96, 0.28, -0.04, and 3.72. The variance of difference between two adjusted means or effects is 2.5600. Note that the intrablock largest and smallest solutions have been shrunk toward zero. This is a characteristic of adjusting for random effects.

4. GAUSS to Recover Interreplicate, Interblock, and Intervariety Information

When conducting a new analysis, it has been useful to write a GAUSS program before attempting to use other packages such as SAS. Accordingly, this was done for the example described in Section 3. The program as well as the output from the program is given in Table A1. Additional statements are easily incorporated into the program to find other sums of squares, e.g., **YC'*YC** / r is the sum of squares for check (ignoring block effects), **YB'*YB** / k is the sum of squares for block (ignoring other effects), and **YN'*YN** minus the correction for the sum of the totals in **YN** is the sum of squares for new (ignoring block). Note that **YN** has the estimated effect μ subtracted from the response in order to remove the mean effect. If **YN** is corrected as above, then the sum of squares is new (ignoring block and replicate effects).

If the totals in Table 1 are obtained by other means, the following statements in Table A1 may be omitted:
Y[rc + n, 1], **X**[rc + n, 1 + r + rb + c + n], and **tot** = **X'*Y**. The matrix X becomes large for most AEDs used in practice.

The program was named augbibd. The totals for the various quantities are obtained as **X'Y**. The solution for the effects is obtained as **b** = $(\mathbf{X'X} - \mathbf{J0})^{-1}\mathbf{X'Y}$. The solutions

Table 2. ANOVAs for data of Table 1.

Source of variation	d.f.		Sum of squares	Mean square
Total	18		2644	
Correction for mean	1		2005.5556	
Replicate (ignoring treatment)	2		269.4444	
Treatment (eliminating replicate, ignoring block)	9		294.3334	
Check		3	65.3333	
Among new		5	217.8889	
Check vs. new		1	11.1111	
Replicate × treatment	6		74.6667	
Block (eliminating treatment)		3	66.6667	22.2222
Intrablock error		3	8.000	2.6667
Replicate (eliminating treatment)	2		200.0000	100.0000
Treatment (ignoring replicate and block)	9		363.7778	
Check (ignoring block)		3	65.3333	
New (ignoring replicate and block)		5	287.3333	
Check vs. new (ignoring replicate and block)		1	11.1111	

Check responses only

Source of variation	d.f.		Sum of squares	Mean square
Total	12		1540	
Correction for mean	1		1200	
Replicate	2		200.0000	
Check (ignoring block)	3		65.3333	
Check × replicate	6		74.6667	
Block (eliminating check)		3	66.6667	22.2222
Intrablock error		3	8.000	2.6667
Check (eliminating block)	3		64.0000	31.3333
New (eliminating replicate and block)	5		111.0794	22.2397
Check vs new (eliminating block and replicate)	1		58.2539	

are identically those used to construct the data set in Table 1 and the responses in **Y**. The solutions for treatment effects eliminating replicate but ignoring block effects, **ter**, will contain block effects. For example, the solution for treatment 1 is 14 = 10 + 4 and the solution for check treatment 7 is -3.6667 = -4 + (4 + 0 -3) / 3. Eliminating the replicate and block effects, the solutions for new, **nbc**, and check, **cbn**, are those used to construct the example.

The symbols EV, EB, RV, and NV in the program are the ANOVA solutions of the error variance component, the block variance component, the replicate variance component, and the new treatment variance component, respectively. The sums of squares for check (eliminating block effect), **css**, and new treatment, (eliminating replicate, and block effects), **nss**, are the Type III sums of squares listed in SAS PROC GLM outputs. Commands for additional sums of squares are easily programmed into the program in Table A1 should they be desired.

5. SAS PROC GLM Program

A simple program for obtaining some of the results for the above example is the following:

```
data augbibd ;
   infile 'augbibd.dat' ;
   input yield rep block treat ;
proc glm data = augbibd ;
   class rep block treat ;
   model yield = rep block(rep) treat / solution ;
  random rep block(rep) ;
  lsmeans treat ;
run ;
```

The data set is named augbibd.dat. Replicate was shortened to rep and treatment to treat as a class variable can have eight or less characters in its name. In the model statement, the term "block(rep) means that blocks are nested within replicates and the term "/ solution" is used to obtain solutions for various effects and may be omitted if only lsmeans are desired. Note that the solutions are obtained by setting the highest numbered effect equal to zero rather than using the constraint that the sum of the effects is zero. Running the above program, we obtain the results in Table A2. Here we note that all sums of squares agree with those from GAUSS. The solutions for replicates are not what is desired but the block and treatment effects are those used to construct the example when the highest numbered effect is added to them. The least squares means, lsmeans, are those obtained from

GAUSS. The ANOVA solutions for the replicate and block variance components are those one would obtain from an analysis of check responses only. By including the random statement for rep and block(rep), the expected values of the replicate and of the blocks (eliminating treatment effects) mean squares are obtained. This is the only use of the random statement in PROC GLM. This program treats both new and check in the same manner as fixed effects. The output for this program is given in Table A2. Annotations for the outputs of programs appear in italics in the Tables.

It may be desirable to obtain an analysis for check responses only. This may be accomplished with the following PROC GLM code:

```
data augbibd ;
   infile 'augbibd.dat' ;
   input yield rep block treat ;
   If treat > 6 and treat < 11 then check = treat ;
proc glm data = augbibd ;
   class rep block check ;
   model yield = rep block(rep) check ;
   random rep block(rep) ;
   lsmeans check;
run ;
```

The output from this program is given in Table A3.

The following PROC GLM program demonstrates how to partition the treatment variable into check and new, to arrange the lsmeans in descending order, and to obtain the ANOVA expected value for replicate and block mean squares.

```
data augbibd ;
  infile 'augbibd.dat ;
  input yield rep block treat ;
proc glm data = augbibd ;
  class rep block treat ;
  model yield = rep block(rep) treat ;
  random rep block(rep) ;
  lsmeans treat / out = lsmeans noprint ;
run;
proc sort data = lsmeans ;
  by descending lsmean ;
proc print ;
run ;
```

```
run ;
```

Additional sums of squares may be obtained from the following program:

```
data augbibd ;
  infile 'augbibd.dat' ;
  input yield rep block treat ;
  if (treat > 6) then new = 0 ; else new = 1 ;
  if (new) then trtn = 999; else trtn = treat ;
proc glm data = augbibd ;
  class rep block treat trtn ;
  model yield = rep block(rep) trtn treat*new ;
  random rep block(rep) ;
  lsmeans trtn ;
run ;
```

The outputs from these programs are given in Table A4.

6. SAS PROC MIXED Program For Recovering Interreplicate, Interblock, and Intervariety Information

When it is desired to recover interreplicate, interblock, and intervariety information, we may use the following PROC MIXED program:

```
data augbibd ;
  infile 'augbibd.dat' ;
  input yield rep block treat ;
  if (treatment > 6) then new = 0 ; else new = 1 ;
  if (new) then trtn = 999 ; else trtn = treat ;
proc mixed data = augbibd info;
  class rep block treat trtn;
  model yield = trtn / solution;
  random rep block(rep) treat*new / solution;
  lsmeans trtn;
  make 'solutionr' out = sr noprint;
run;
proc sort data = sr;
  by descending est;
proc print;
```

```
run;
```

REML solutions for the variance components are used in the SAS package. A new variable trtn is created to divide the treatments into check and new treatments. The solutions sr are created and then placed in descending order. This is a very useful device for screening experiments as all the new treatments are ranked and the experimenter may easily observe which are the top performers and which are low performers. All check treatments in sr are given the value of zero. Then it is easy to observe those out-performing the checks.

If it is desired to use solutions close to ANOVA solutions we may change the PROC MIXED statement to read:

```
proc mixed data = augibib.dat info itdetails maxiter=1 method = reml;
```

Maxiter = 1 denotes that only one iteration will be used to obtain the variance component estimates. Method = reml denotes that REML solutions are desired. Other methods such as ml, minque, and mivque may be used if desired. The output for this program at iteration one for REML solutions, which are approximately ANOVA solutions, is given in Table A6. The output for the full number of iterations is given in Table A5. These REML solutions were obtained using the bounds in the SAS package. If it is desired to use no constraints on the REML solutions, simply add the words "nobounds" in the PROC MIXED statement just before the semicolon as follows:

```
prtoc mixed data = augbibd.dat info nobounds;
```

The output for this program appears in Table A7. The effect of the various codes used above on the solutions for variance components may be noted from the outputs. The above also denotes the flexibility one has using the SAS package.

7. Some Comments

The above codes for recovering the various types of information desired is easily extended to other experiment designs. For the augmented row-column designs as given by Federer *et al.* (1975, 1975) and the example described in Federer (1996), we may use the following program. Let the file name be augrc1.dat for the 15 row by 12 column design with c = 2 check varieties each included 30 times and n = 120 new genotypes each in once. Since the design was not connected, fourth degree polynomial contrasts of rows (R1, R2, R3, R4) and of columns (C1, C2, C3, C4) plus the row linear by column linear (R1 × C1) and row linear by column quadratic (R1 × C2) interactions were used for the analysis. The following code was used for this data set at site 1:

```
data augrc1 ;
   infile 'augrc1.dat' ;
   input site row col treat GW C1 C2 C3 C4 R1 R2 R3 R4 ;
   if (treat > 120) then new = 0; else new = 1 ;
   if (new) then trtn = 999; else trtn = treat ;
   LL = R1*C1 ; LQ = R1*C2 ;
proc mixed data = augrc1 info ;
   class row col treat trtn ;
   model yield = R1 R2 R3 R4 C1 C2 C3 C4 LL LQ treat*new /solution ;
   random row col treat*new / solution ;
   lsmeans trtn ;
   make 'solutionr' out = sr noprint ;
run ;
proc sort ;
   by descending est ;
proc print ;
run ;
```

The Ri and Ci (i = 1, 2, 3, 4) contrasts are row and column contrasts and row and column effects are random. The new treatment solutions have been adjusted for interrow, intercolumn, and intervariety information. The above results may be extended for any augmented design.

8. Literature

Cochran, W. G. and G. M. Cox (1957). *Experimental Designs*, second edition. John Wiley & Sons, Inc., New York.

Federer, W. T. (1955). *Experimental Design--Theory and Application.* Macmillan Co., New York (Republished by Oxford & IBH Publishing Co., New Delhi, 1967 and 1974).

Federer, W. T. (1956a). A method for evaluating genetic progress in a sugar cane breeding program. Hawaiian Planters' Record LV(2):177-189.

Federer, W. T. (1956b). Augmented (or hoonuiaku) designs. Hawaiian Planters' Record LV(2):191-208.

Federer, W. T. (1961). Augmented designs with one-way elimination of heterogeneity. Biometrics 17:447-473.

Federer, W. T. (1991). Statistics and Society. Marcel Dekker, Inc. New York, Section 7.11.

Federer, W. T. (1996). Recovery of interblock, intergradient, and intervariety information in incomplete block and lattice rectangle designed experiments. Biometrics (in press)

Federer, W. T., R. C. Nair, and D. Raghavarao (1975). Some augmented row-column designs. Biometrics 31:361-373.

Federer, W. T. and D. Raghavarao (1975). On augmented designs. Biometrics 31:29-35.

SAS Institute Inc. (1989). SAS/STAT Users Guide, Version 6, Fourth Edition, Volume 2. SAS Institute Inc., Cary N. C.

SAS Institute Inc. (1996). SAS/STAT Software: Changes and Enhancements Through Release 6.11. SAS Institute Inc., Cary N.C.

Yates, F. (1936). A new method of arranging variety trials involving a large number of varieties. Journal of Agricultural Science 26:424-455.

```
output file = c:\gauss\table_a1.out reset;
@Table A1.  Gauss program and output for analyzing an augmented incomplete block
 design with recovery of interblock and intervariety information.@
"The rc + n vector of observations (yield) are";
let Y[18,1] = 4 10 19 2 4 1 7 9 10 9 15 5 8 16 19 18 18 16;
"The design matrix X[rc+n,1+r+rb+c+n] is"; let X[18,20] =
1 1 0 0  1 0 0 0 0 0  0 0 0 0 0 0  1 0 0 0
1 1 0 0  1 0 0 0 0 0  0 0 0 0 0 0  0 1 0 0
1 1 0 0  1 0 0 0 0 0  1 0 0 0 0 0  0 0 0 0
1 1 0 0  0 1 0 0 0 0  0 0 0 0 0 0  0 0 1 0
1 1 0 0  0 1 0 0 0 0  0 0 0 0 0 0  0 0 0 1
1 1 0 0  0 1 0 0 0 0  0 1 0 0 0 0  0 0 0 0
1 0 1 0  0 0 1 0 0 0  0 0 0 0 0 0  1 0 0 0
1 0 1 0  0 0 1 0 0 0  0 0 0 0 0 0  0 0 1 0
1 0 1 0  0 0 1 0 0 0  0 0 1 0 0 0  0 0 0 0
1 0 1 0  0 0 0 1 0 0  0 0 0 0 0 0  0 1 0 0
1 0 1 0  0 0 0 1 0 0  0 0 0 0 0 0  0 0 0 1
1 0 1 0  0 0 0 1 0 0  0 0 0 1 0 0  0 0 0 0
1 0 0 1  0 0 0 0 1 0  0 0 0 0 0 0  1 0 0 0
1 0 0 1  0 0 0 0 1 0  0 0 0 0 0 0  0 0 0 1
1 0 0 1  0 0 0 0 1 0  0 0 0 0 1 0  0 0 0 0
1 0 0 1  0 0 0 0 0 1  0 0 0 0 0 0  0 1 0 0
1 0 0 1  0 0 0 0 0 1  0 0 0 0 0 0  0 0 1 0
1 0 0 1  0 0 0 0 0 1  0 0 0 0 0 1  0 0 0 0;
"The various totals are obtained as";
format 2,3; tot = X'*Y; tot';
Z = zeros(10,20); Z6 = zeros(6,6); Z10 = zeros(6,10); Z46 = zeros(4,6);
let Z4[4,4] = 0 6 6 6  0 0 0 0  0 0 0 0  0 0 0 0;  J1 = ones(6,4);
let J4[4,4] = 3 3 3 3  1 1 1 1  1 1 1 1  1 1 1 1;
let J6[4,6] = 3 3 3 3 3 3  3 3 0 0 0 0  0 0 3 3 0 0  0 0 0 0 3 3;
M1 = Z4~J6~Z46~J4; M2 = Z10~Z6~J1; J0 = M1|Z|M2; XC = X'*X - J0;
b = inv(XC)*X'*Y; b';
"Block totals minus s = 3 times mean plus replicate effect are";
let YB[6,1] = 18 -8 -4 -1 -2 7;
"Check treatment totals minus r = 3 times mean effect are";
let YC[4,1] = -11 7 -1 5;
"New treatment totals minus mean plus replicate effect are";
let YN[6,1] = 14 -4 0 -5 4 1;
"New treatment totals minus mean effect are";
let YNm[6,1] = 9 -9 0 -5 9 6;
"Replicate totals minus c + n/3 = 6 times mean effect
and minus c + n/3 =  6 times overall average are";
let YR[3,1] = -20 -5 35; let YRm[3,1] = -23.3333 -8.3333 31.6667;
"Replicate sum of squares ignoring all other effects except mean is";
format 2,8; YRm'*YRm/6;
let C[4,6] = 1 0 1 0 1 0  1 0 0 1 0 1  0 1 1 0 0 1  0 1 0 1 1 0;
I4 = eye(4); I6 = eye(6); NB = I6;
"Total sum of squares is Y'*Y"; tss = Y'*Y; tss;
let RT[3,10] =
1 1 0 0 0 0 0 0 0 0
0 0 1 1 0 0 0 0 0 0
0 0 0 0 1 1 0 0 0 0;
U = (I6~Z46'|Z46~(3*I4)) - (RT'*RT/6);
V = (YNm|YC) - RT'*YR/6;
```

```
"Treatment effects adjusted for replicate effects ter are";
ter = inv(U)*V; ter';
"Treatment (eliminating replicate effect) sum of squares tss is";
tss = ter'*V; tss; s = 3; r = 3; k = 2;
W = 3*I6 - (NB'~(C' - J1))*inv((I6~Z46')|(Z46~(r*I4)))*(NB|C);
S1 = YB - (NB'~(C' - J1))*inv((I6~Z46')|(Z46~(r*I4)))*(YN|YC);
"Block Type III effects ben and sum of squares bss are";
ben = inv(W)*S1;  ben'; bss = ben'*S1; bss;
P = r*I4 - ((C - J1')~Z46)*inv(((s*I6)~NB)|(NB'~I6))*(C'|Z46');
Q = YC - ((C - J1')~Z46)*inv(((s*I6)~NB)|(NB'~I6))*(YB|YN);
"Check Type III effects cbn and sum of squares css are";
cbn = inv(P)*Q; cbn'; css = cbn'*Q; css;
M = I6 - (NB~Z46')*inv(((s*I6)~(C' - J1))|(C~(r*I4)))*(NB'|Z46);
O = YN - (NB~Z46')*inv(((s*I6)~(C' - J1))|(C~(r*I4)))*(YB|YC);
"New treatment Type III effects nbc and sum of squares css are";
nbc = inv(M)* O; nbc'; nss = nbc'*O; nss;
EV=2.6667; BV = 14.6667; NV = 19.5492; RV = 17.0000;
Pa= r*I4 -((C-J1')~Z46)*inv((((s + EV/BV)*I6)~NB)|(NB'~I6))*(C'|Z46');
Qa = YC - ((C - J1')~Z46)*inv((((s + EV/BV)*I6)~NB)|(NB'~I6))*(YB|YN);
"Check effects cbna with recovery of interblock information are";
cbna = inv(Pa)*Qa; cbna';
"Variance-covariance matrix for cbna is inv(Pa)*EV";
varch = EV*inv(Pa); varch;
ZC = zeros(3,4); ZB = zeros(3,6); I3 = eye(3);
"Block totals minus s times mean effect"; let YBm = 3 -23 -4 -1 13 22;
"Incidence matrix for replicates and new treatments";
let RN[3,6] = 1 1 0 0 0 0  0 0 1 1 0 0  0 0 0 0 1 1;
Ma = (1 + EV/NV)*I6 - (RN'~NB'~Z46')*inv((((6 + EV/RV)*I3)~ZB~ZC)|
(ZB'~((s + EV/BV)*I6)~(C' - J1))|(ZC'~C~(r*I4)))*(RN|NB|Z46);
Oa = YNm - (RN'~NB'~Z46')*inv((((6 + EV/RV)*I3)~ZB~ZC)|
(ZB'~((s + EV/BV)*I6)~(C'-J1))|(ZC'~C~(r*I4)))*(YR|YB|YC);
"New treatments effects nrbc with recovery of interreplicate, interblock,
and intervariety information and variance-covariance matrix";
nrbc = inv(Ma)*Oa; nrbc';
varnew = EV*inv(Ma);varnew;
output file = c:\gauss\table_a1.out off;
```

The rc + n vector of observations (yield) are
The design matrix X[rc+n,1+r+rb+c+n] is
The various totals are obtained as
190. 40.0 55.0 95.0 33.0 7.00 26.0 29.0 43.0 52.0 19.0 1.00 10.0 5.00 19.0
16.0 19.0 37.0 29.0 35.0
10.0 -5.00 1.11e-15 5.00 4.00 -4.00 -1.35e-15 -4.06e-16 -3.00 3.00 10.0
-1.13e-15 -4.79e-16 -5.00 7.00 -2.00 -4.00 1.28e-15 1.24e-15 4.00
Block totals minus s = 3 times mean plus replicate effect are
Check treatment totals minus r = 3 times mean effect are
New treatment totals minus mean plus replicate effect are
New treatment totals minus mean effect are
Replicate totals minus c + n/3 = 6 times mean effect
and minus c + n/3 = 6 times overall average are
Replicate sum of squares ignoring all other effects except mean is
269.44444
Total sum of squares is Y'*Y
2644.0000
Treatment effects adjusted for replicate effects ter are
14.000000 -4.0000000 8.7928796e-17 -5.0000000 4.0000000 1.0000000 -3.6666667
2.3333333 -0.33333333 1.6666667
Treatment (eliminating replicate effect) sum of squares tss is
294.33333
Block Type III effects ben and sum of squares bss are
4.0000000 -4.0000000 -1.0176593e-16 9.2509550e-17 -3.0000000 3.0000000
66.666667
Check Type III effects cbn and sum of squares css are
-4.0000000 2.0355896e-16 9.2536655e-17 4.0000000
64.000000
New treatment Type III effects nbc and sum of squares css are
10.000000 -9.7698474e-16 -2.4721504e-17 -5.0000000 7.0000000 -2.0000000
111.07937
Check effects cbna with recovery of interblock information are
-3.9599996 0.28000252 -0.040000360 3.7199975
Variance-covariance matrix for cbna is inv(Pa)*EV

1.0759552 -0.20406035 -0.20406035 -0.20406035
-0.20406035 1.0759552 -0.20406035 -0.20406035
-0.20406035 -0.20406035 1.0759552 -0.20406035
-0.20406035 -0.20406035 -0.20406035 1.0759552
Block totals minus s times mean effect
Incidence matrix for replicates and new treatments
New treatments effects nrbc with recovery of interreplicate, interblock,
and intervariety information and variance-covariance matrix
7.9301526 -1.0062887 0.31384673 -3.7629937 5.5356713 -1.2155712

4.3107074 0.70130130 -0.17526307 -0.17526307 -0.17526307 -0.17526307
0.70130130 4.3107074 -0.17526307 -0.17526307 -0.17526307 -0.17526307
-0.17526307 -0.17526307 4.3107074 0.70130130 -0.17526307 -0.17526307
-0.17526307 -0.17526307 0.70130130 4.3107074 -0.17526307 -0.17526307
-0.17526307 -0.17526307 -0.17526307 -0.17526307 4.3107074 0.70130130
-0.17526307 -0.17526307 -0.17526307 -0.17526307 0.70130130 4.3107074

Table A2. Annotated GLM output and program for an incomplete block design with v = 10 treatments in incomplete blocks of size s = 3 and r = 3 replicates.

Class Level Information

Class	Levels	Values
REP	3	1 2 3
BLOCK	2	1 2
TREAT	10	1 2 3 4 5 6 7 8 9 10

Number of observations in data set = 18

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	630.444444	45.031746	16.89	0.0197
Error	3	8.000000	2.666667		
Corrected Total	17	638.444444			

R-Square	C.V.	Root MSE	YIELD Mean
0.987470	15.47046	1.63299	10.5556

Source	DF	Type I SS	Mean Square	F Value	Pr > F
REP	2	269.444444	134.722222	50.52	0.0049
BLOCK(REP)	3	127.666667	42.555556	15.96	0.0239
TREAT	9	233.333333	25.925926	9.72	0.0437

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	200.000000	100.000000	37.50	0.0075
BLOCK(REP)	3	66.666667	22.222222	8.33	0.0576
TREAT	9	233.333333	25.925926	9.72	0.0437

Parameter		Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
INTERCEPT		22.00000000 B	12.05	0.0012	1.82574186
REP	1	-17.00000000 B	-9.31	0.0026	1.82574186
	2	-8.00000000 B	-4.38	0.0220	1.82574186
	3	0.00000000 B	.	.	.
BLOCK(REP)	1 1	8.00000000 B	4.00	0.0280	2.00000000
	2 1	0.00000000 B	.	.	.
	1 2	-0.00000000 B	-0.00	1.0000	2.00000000
	2 2	0.00000000 B	.	.	.
	1 3	-6.00000000 B	-3.00	0.0577	2.00000000
BLOCK(REP)	2 3	0.00000000 B	.	.	.
TREAT	1	6.00000000 B	2.45	0.0917	2.44948974
	2	-4.00000000 B	-1.85	0.1612	2.16024690
	3	-4.00000000 B	-1.63	0.2010	2.44948974
	4	-9.00000000 B	-4.17	0.0252	2.16024690
	5	3.00000000 B	1.39	0.2591	2.16024690
	6	-6.00000000 B	-2.45	0.0917	2.44948974
	7	-8.00000000 B	-4.90	0.0163	1.63299316
	8	-4.00000000 B	-2.45	0.0917	1.63299316
	9	-4.00000000 B	-2.45	0.0917	1.63299316
	10	0.00000000 B	.	.	.

Note : To obtain solutions, the highest numbered effect is set equal to zero in SAS PROC GLM rather than using the constraint that the sum of the effects equals zero. This means all the above solutions have the effect that is set equal to zero subtracted from the other effects in the variable. The standard errors of estimate are standard errors of a difference between two effects.

NOTE: The X'X matrix has been found to be singular and a generalized inverse was used to solve the normal equations. Estimates followed by the letter 'B' are biased, and are not unique estimators of the parameters.

Source Type III Expected Mean Square

REP Var(Error) + 2 Var(BLOCK(REP)) + 4 Var(REP)

BLOCK(REP) Var(Error) + 1.3333 Var(BLOCK(REP))

TREAT Var(Error) + Q(TREAT)

These expectations of mean squares are those one would obtain for check treatments only. This is becuse there is only experimental unit for each new treatment which goes to obtaining a solution for the new treatment effects and nothing to estimating block effects, replicate effects, or the overall mean effect.

The coefficient $1.3333 = k(r-1)/r = 2(3-1)/3$. *The block size for checks is* $k = 2$.

Least Squares Means

TREAT	YIELD LSMEAN
1	20.0000000
2	10.0000000
3	10.0000000
4	5.0000000
5	17.0000000
6	8.0000000
7	6.0000000
8	10.0000000
9	10.0000000
10	14.0000000

These are the correct least squares means using the values for constructing the data set. For example, $10 + 10 =$ *mean effect plus treatment one effect* $= 20$.

PROC GLM program

```
options ls = 76;               (This sets the line length and ps = nn sets page size.)
data augbibd;                  (Name of data set or file to be used.)
  infile 'augbibd.sas';        (Actual filename of data set.)
  input yield rep block treat;(These are the column identifiers of the data.)
proc glm data = augbibd;       (The procedure used is proc glm.)
  class rep block treat;       (Class variables are discrete valued variables.)
  model yield = rep block(rep) treat / solution;(This is the usual linear model.)
  random rep block(rep);       (Class and model entries only. This statement
                               is included only if expectations of mean squares
                               are desired.)
  lsmeans treat;               (Intrablock (fixed effect) treatment means.)
run;                           ("/ solution" or "/s" will give solutions for all
                               variables in the model statement.)
```

Table A3. Annotated output and GLM program for data set augbibd, check responses only.

Class Level Information

Class	Levels	Values
REP	3	1 2 3
BLOCK	2	1 2
CHECK	4	7 8 9 10

Number of observations in data set = 18

NOTE: Due to missing values, only 12 observations can be used in this analysis.

Dependent Variable: YIELD

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	332.000000	41.500000	15.56	0.0227
Error	3	8.000000	2.666667		
Corrected Total	11	340.000000			

R-Square	C.V.	Root MSE	YIELD Mean
0.976471	16.32993	1.63299	10.0000

Dependent Variable: YIELD

Source	DF	Type I SS	Mean Square	F Value	Pr > F
REP	2	200.000000	100.000000	37.50	0.0075
BLOCK(REP)	3	68.000000	22.666667	8.50	0.0561
CHECK	3	64.000000	21.333333	8.00	0.0607

Source	DF	Type III SS	Mean Square	F Value	Pr > F
REP	2	200.000000	100.000000	37.50	0.0075
BLOCK(REP)	3	66.666667	22.222222	8.33	0.0576

CHECK	3	64.000000	21.333333	8.00	0.0607

Parameter		Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
INTERCEPT		22.00000000 B	12.05	0.0012	1.82574186
REP	1	-17.00000000 B	-9.31	0.0026	1.82574186
	2	-8.00000000 B	-4.38	0.0220	1.82574186
	3	0.00000000 B	.	.	.
BLOCK(REP)	1 1	8.00000000 B	4.00	0.0280	2.00000000
	2 1	0.00000000 B	.	.	.
	1 2	-0.00000000 B	-0.00	1.0000	2.00000000
	2 2	0.00000000 B	.	.	.
	1 3	-6.00000000 B	-3.00	0.0577	2.00000000

Parameter		Estimate	T for H0: Parameter=0	Pr > |T|	Std Error of Estimate
BLOCK(REP)	2 3	0.00000000 B	.	.	.
CHECK	7	-8.00000000 B	-4.90	0.0163	1.63299316
	8	-4.00000000 B	-2.45	0.0917	1.63299316
	9	-4.00000000 B	-2.45	0.0917	1.63299316
	10	0.00000000 B	.	.	.

NOTE: The X'X matrix has been found to be singular and a generalized inverse was used to solve the normal equations. Estimates followed by the letter 'B' are biased, and are not unique estimators of the parameters.

Source	Type III Expected Mean Square
REP	Var(Error) + 2 Var(BLOCK(REP)) + 4 Var(REP)
BLOCK(REP)	Var(Error) + 1.3333 Var(BLOCK(REP))
CHECK	Var(Error) + Q(CHECK)

Least Squares Means

CHECK	YIELD LSMEAN	
7	6.0000000	*(These are the intrablock (fixed*
8	10.0000000	*effect) check means and are those*
9	10.0000000	*used to construct the data. For*
10	14.0000000	*example,* $10 - 4 = 6, 10 + 0 =$

$10, 10+0 = 10, 10+4 = 14.)$

PROC GLM Program (For check reponses only.)

```
options ls = 76;
data augbibd;
  infile 'augbibd.sas';
  input yield rep block treat;
  if treat > 6 and treat < 11 then check = treat; (Divides into checks and new.)
proc glm data = augbibd;
  class rep block check;
  model yield = rep block(rep) check / solution;(Solutions for model effects.)
  random rep block(rep);   lsmeans check;
run;
```

Table A4. Annotated output and PROC GLM program for ordering lsmeans.

Class Level Information

Class	Levels	Values
R	3	1 2 3
B	2	1 2
TR	10	1 2 3 4 5 6 7 8 9 10
TRTN	5	7 8 9 10 999

Number of observations in data set = 18

Dependent Variable: Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	630.4444444	45.0317460	16.89	0.0197
Error	3	8.0000000	2.6666667		
Corrected Total	17	638.4444444			

R-Square	C.V.	Root MSE	Y Mean
0.987470	15.47046	1.632993	10.55556

Dependent Variable: Y

Source	DF	Type I SS	Mean Square	F Value	Pr > F
R	2	269.4444444	134.7222222	50.52	0.0049
B(R)	3	127.6666667	42.5555556	15.96	0.0239
TR	9	233.3333333	25.9259259	9.72	0.0437

Source	DF	Type III SS	Mean Square	F Value	Pr > F
R	2	200.0000000	100.0000000	37.50	0.0075
B(R)	3	66.6666667	22.2222222	8.33	0.0576
TR	9	233.3333333	25.9259259	9.72	0.0437

Source	Type III Expected Mean Square
R	Var(Error) + 2 Var(B(R)) + 4 Var(R)
B(R)	Var(Error) + 1.3333 Var(B(R))
TR	Var(Error) + Q(TR)

OBS	_NAME_	TR	LSMEAN	STDERR
1	Y	1	20	2.02759
2	Y	5	17	2.02759
3	Y	10	14	1.10554
4	Y	3	10	2.02759
5	Y	8	10	1.10554
6	Y	9	10	1.10554
7	Y	2	10	2.02759
8	Y	6	8	2.02759
9	Y	7	6	1.10554
10	Y	4	5	2.02759

(*Intrablock, fixed effects, least squares means arranged in descending order.*)

(*PROC GLM code for obtaining least squares means in descending order.*)

```
data augbibd;
  infile 'augbibd.sas';
  input y r b tr;
proc glm data = augbibd;
  class r b tr trtn;
  model y = r b(r) tr;
  random r b(r);
  lsmeans tr/out = lsmeans   noprint;
run;
proc sort data = lsmeans;
  by descending lsmean;
proc print;
run;
```

Table A5. Annotated PROC MIXED output and program for augbibd data set.

Model Information

Description	Value
Data Set	WORK.AUGBIBD
Dependent Variable	YIELD
Covariance Structure	Variance Components
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment
Estimation Method	REML (*Restricted maximum likelihood.*)

Class Level Information

Class	Levels	Values
REP	3	1 2 3
BLOCK	2	1 2
TREAT	10	1 2 3 4 5 6 7 8 9 10
TRTN	5	7 8 9 10 999

(999 *created to divide treatments into two sets, one fixed and the other, new, random.*)

Dimensions

Description	Value
Covariance Parameters	4
Columns in X	6

Dimensions

Description	Value
Columns in Z	19
Subjects	1
Max Obs Per Subject	18
Observations Used	18
Observations Not Used	0
Total Observations	18

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	68.15139103	
1	2	58.12205688	0.00113311
2	1	58.08507883	0.00005579
3	1	58.08339618	0.00000023
4	1	58.08338961	0.00000000

(Note the large change in evaluations column from iteration 0 to iteration 1.)

Convergence criteria met.

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > |Z|
REP	6.24074288	16.16922515	25.52889581	0.63	0.5265
BLOCK(REP)	5.96902124	15.46521787	14.15284025	1.09	0.2745
NEW*TREAT	11.01774134	28.54601508	20.87830356	1.37	0.1715
Residual	1.00000000	2.59091353	2.04566328	1.27	0.2053

(The ANOVA solution for the residual variance component is $2.66667 = 8/3$.)

Model Fitting Information for YIELD

Description	Value
Observations	18.0000
Variance Estimate	2.5909

Model Fitting Information for YIELD

Description	Value
Standard Deviation Estimate	1.6096
REML Log Likelihood	-40.9879
Akaike's Information Criterion	-44.9879
Schwarz's Bayesian Criterion	-46.1178
-2 REML Log Likelihood	81.9758

Solution for Fixed Effects

Parameter	Estimate	Std Error	DDF	T	Pr > |T|
INTERCEPT	11.66666667	3.62722568	2	3.22	0.0846
TRTN 7	-5.84153043	2.51509959	3	-2.32	0.1028
TRTN 8	-1.38487404	2.51509959	3	-0.55	0.6202
TRTN 9	-1.57935168	2.51509959	3	-0.63	0.5746
TRTN 10	2.13908949	2.51509959	3	0.85	0.4575
TRTN 999	0.00000000				

(The check treatment effects have been adjusted for interblock information.)

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
TRTN	4	3	6.65	0.0758

Least Squares Means

Level	LSMEAN	Std Error	DDF	T	Pr > |T|
TRTN 7	5.82513624	3.01720332	3	1.93	0.1491
TRTN 8	10.28179262	3.01720332	3	3.41	0.0422
TRTN 9	10.08731498	3.01720332	3	3.34	0.0443
TRTN 10	13.80575615	3.01720332	3	4.58	0.0196
TRTN 999	11.66666667	3.62722568	3	3.22	0.0487

(These are the check means adjusted for interblock information.)

OBS	PARM	EST	SE_PRED	DDF	T	P_T
1	NEW*TREAT 1	7.55175405	2.78641001	3	2.71	0.0732
2	NEW*TREAT 5	4.46844822	2.78641001	3	1.60	0.2071
3	BLOCK(REP) 2 3	4.03339310	2.63153067	3	1.53	0.2229
4	REP 3	3.31860074	3.01031216	3	1.10	0.3508
5	BLOCK(REP) 1 1	2.30502968	2.63153067	3	0.88	0.4455
6	BLOCK(REP) 1 2	0.07884568	2.63153067	3	0.03	0.9780
7	NEW*TREAT 7	0.00000000	5.34284709	3	0.00	1.0000
8	NEW*TREAT 8	0.00000000	5.34284709	3	0.00	1.0000
9	NEW*TREAT 9	0.00000000	5.34284709	3	0.00	1.0000
10	NEW*TREAT 10	0.00000000	5.34284709	3	0.00	1.0000
11	REP 2	-0.10973273	3.01031216	3	-0.04	0.9732
12	BLOCK(REP) 2 2	-0.18380065	2.63153067	3	-0.07	0.9487
13	BLOCK(REP) 1 3	-0.85928408	2.63153067	3	-0.33	0.7655
14	NEW*TREAT 3	-1.49966588	2.78641001	3	-0.54	0.6278
15	NEW*TREAT 2	-1.91023667	2.78641001	3	-0.69	0.5422
16	NEW*TREAT 6	-2.76747683	2.78641001	3	-0.99	0.3938
17	REP 1	-3.20886801	3.01031216	3	-1.07	0.3646
18	BLOCK(REP) 2 1	-5.37418373	2.63153067	3	-2.04	0.1338
19	NEW*TREAT 4	-5.84282288	2.78641001	3	-2.10	0.1269

(The solutions for all random effects have been ordered from largest to smallest. Note that the check treatment effects are zero. The new effects have been adjusted for interreplicate, interblock, and intervariety (new) information.))

```
options ls = 76;
data augbibd;
  infile 'augbibd.sas';
  input yield rep block treat;
  if (treat > 6) then new = 0; else new = 1;
  if (new) then trtn = 999; else trtn = treat;
proc mixed data = augbibd info ;
  class rep block treat trtn;
  model yield = trtn / solution;
  random rep block(rep) treat*new / solution;
  lsmeans trtn;
  make 'solutionr' out = sr noprint;
run;
proc sort data = sr;
  by descending est;
proc print;
run;
```

Table A6. Annotated output and program for one iteration of REML.

Model Information

Description	Value
Data Set	WORK.AUGBIBD
Dependent Variable	YIELD
Covariance Structure	Variance Components
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment
Estimation Method	REML

Class Level Information

Class	Levels	Values
REP	3	1 2 3
BLOCK	2	1 2
TREAT	10	1 2 3 4 5 6 7 8 9 10
TRTN	5	7 8 9 10 999

Dimensions

Description	Value
Covariance Parameters	4
Columns in X	6

Description	Value
Columns in Z	19
Subjects	1
Max Obs Per Subject	18
Observations Used	18
Observations Not Used	0
Total Observations	18

REML Estimation Iteration History

REP	BLOCK(REP)	NEW*TREAT	Residual	Iteration
0.00000000	0.00000000	0.00000000	43.23076923	0
14.18954804	15.13827363	26.83449996	2.96084596	1

Since the number of iterations was limited to one, the iteration did not converge.

Evaluations	Objective	Criterion
1	68.15139103	
2	58.12205688	0.00113311

Did not converge.

Covariance Parameters at
Last REML Iteration

Cov Parm	Ratio	Estimate
REP	4.79239657	14.18954804
BLOCK(REP)	5.11282039	15.13827363
NEW*TREAT	9.06311922	26.83449996
Residual	1.00000000	2.96084596

(*The ANOVA solution was* 2.6667.)

OBS	PARM	EST	SE PRED	DDF	T	P_T
1	NEW*TREAT 1	7.55175405	2.78641001	3	2.71	0.0732
2	NEW*TREAT 5	4.46844822	2.78641001	3	1.60	0.2071
3	BLOCK(REP) 2 3	4.03339310	2.63153067	3	1.53	0.2229
4	REP 3	3.31860074	3.01031216	3	1.10	0.3508
5	BLOCK(REP) 1 1	2.30502968	2.63153067	3	0.88	0.4455
6	BLOCK(REP) 1 2	0.07884568	2.63153067	3	0.03	0.9780
7	NEW*TREAT 7	0.00000000	5.34284709	3	0.00	1.0000
8	NEW*TREAT 8	0.00000000	5.34284709	3	0.00	1.0000
9	NEW*TREAT 9	0.00000000	5.34284709	3	0.00	1.0000
10	NEW*TREAT 10	0.00000000	5.34284709	3	0.00	1.0000
11	REP 2	-0.10973273	3.01031216	3	-0.04	0.9732
12	BLOCK(REP) 2 2	-0.18380065	2.63153067	3	-0.07	0.9487
13	BLOCK(REP) 1 3	-0.85928408	2.63153067	3	-0.33	0.7655
14	NEW*TREAT 3	-1.49966588	2.78641001	3	-0.54	0.6278
15	NEW*TREAT 2	-1.91023667	2.78641001	3	-0.69	0.5422
16	NEW*TREAT 6	-2.76747683	2.78641001	3	-0.99	0.3938
17	REP 1	-3.20886801	3.01031216	3	-1.07	0.3646
18	BLOCK(REP) 2 1	-5.37418373	2.63153067	3	-2.04	0.1338
19	NEW*TREAT 4	-5.84282288	2.78641001	3	-2.10	0.1269

```
options ls = 76;
data augbibd;
  infile 'augbibd.sas';
  input yield rep block treat;
  if (treat > 6) then new = 0; else new = 1;
  if (new) then trtn = 999; else trtn = treat;
proc mixed data = augbibd info itdetails maxiter=1 method = reml;
  class rep block treat trtn;
  model yield = trtn / solution;
  random rep block(rep) treat*new / solution;
  lsmeans trtn;
  make 'solutionr' out = sr noprint;run;
proc sort data = sr;
  by descending est;
proc print;
run;
```

Table A7. Annotated output and program for PROC MIXED with no bounds on REML solutions.

Model Information

Description	Value
Data Set	WORK.AUGBIBD
Dependent Variable	YIELD
Covariance Structure	Variance Components
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Containment
Estimation Method	REML

Class Level Information

Class	Levels	Values
REP	3	1 2 3
BLOCK	2	1 2
TREAT	10	1 2 3 4 5 6 7 8 9 10
TRTN	5	7 8 9 10 999

Dimensions

Description	Value
Covariance Parameters	4
Columns in X	6
Columns in Z	19
Subjects	1
Max Obs Per Subject	18
Observations Used	18
Observations Not Used	0
Total Observations	18

REML Estimation Iteration History

Iteration	Evaluations	Objective	Criterion
0	1	68.15139103	
1	2	58.12205688	0.00113311
2	1	58.08507883	0.00005573
3	1	58.08339618	0.00000023
4	1	58.08338961	0.00000000

Convergence criteria met.

(With no bounds on the REML solutions for the variance components, only four iterations were required for convergence.)

Covariance Parameter Estimates (REML)

Cov Parm	Ratio	Estimate	Std Error	Z	Pr > |Z|
REP	6.24074288	16.16922515	25.52889581	0.63	0.5265
BLOCK(REP)	5.96902124	15.46521787	14.15284025	1.09	0.2745
NEW*TREAT	11.01774134	28.54601508	20.87830356	1.37	0.1715
Residual	1.00000000	2.59091353	2.04566328	1.27	0.2053

(*The residual variance component was fairly close to the ANOVA solution, i.e.*, 2.6667.)

Model Fitting Information for YIELD

Description	Value
Observations	18.0000
Variance Estimate	2.5909

Model Fitting Information for YIELD

Description	Value
Standard Deviation Estimate	1.6096
REML Log Likelihood	-40.9879
Akaike's Information Criterion	-44.9879
Schwarz's Bayesian Criterion	-46.1178
-2 REML Log Likelihood	81.9758
Null Model LRT Chi-Square	10.0680
Null Model LRT DF	3.0000
Null Model LRT P-Value	0.0180

Solution for Fixed Effects

Parameter	Estimate	Std Error	DDF	T	Pr > |T|
INTERCEPT	11.66666667	3.62722568	2	3.22	0.0846
TRTN 7	-5.84153043	2.51509959	3	-2.32	0.1028
TRTN 8	-1.38487404	2.51509959	3	-0.55	0.6202
TRTN 9	-1.57935168	2.51509959	3	-0.63	0.5746
TRTN 10	2.13908949	2.51509959	3	0.85	0.4575
TRTN 999	0.00000000	.	.	.	.

Tests of Fixed Effects

Source	NDF	DDF	Type III F	Pr > F
TRTN	4	3	6.65	0.0758

Least Squares Means

Level	LSMEAN	Std Error	DDF	T	Pr > |T|
TRTN 7	5.82513624	3.01720332	3	1.93	0.1491
TRTN 8	10.28179262	3.01720332	3	3.41	0.0422
TRTN 9	10.08731498	3.01720332	3	3.34	0.0443
TRTN 10	13.80575615	3.01720332	3	4.58	0.0196
TRTN 999	11.66666667	3.62722568	3	3.22	0.0487

OBS	PARM	EST	SE_PRED	DDF	T	P_T
1	NEW*TREAT 1	7.55175405	2.78641001	3	2.71	0.0732
2	NEW*TREAT 5	4.46844822	2.78641001	3	1.60	0.2071
3	BLOCK(REP) 2 3	4.03339310	2.63153067	3	1.53	0.2229
4	REP 3	3.31860074	3.01031216	3	1.10	0.3508
5	BLOCK(REP) 1 1	2.30502968	2.63153067	3	0.88	0.4455
6	BLOCK(REP) 1 2	0.07884568	2.63153067	3	0.03	0.9780
7	NEW*TREAT 7	0.00000000	5.34284709	3	0.00	1.0000
8	NEW*TREAT 8	0.00000000	5.34284709	3	0.00	1.0000
9	NEW*TREAT 9	0.00000000	5.34284709	3	0.00	1.0000
10	NEW*TREAT 10	0.00000000	5.34284709	3	0.00	1.0000
11	REP 2	-0.10973273	3.01031216	3	-0.04	0.9732
12	BLOCK(REP) 2 2	-0.18380065	2.63153067	3	-0.07	0.9487
13	BLOCK(REP) 1 3	-0.85928408	2.63153067	3	-0.33	0.7655
14	NEW*TREAT 3	-1.49966588	2.78641001	3	-0.54	0.6278
15	NEW*TREAT 2	-1.91023667	2.78641001	3	-0.69	0.5422
16	NEW*TREAT 6	-2.76747683	2.78641001	3	-0.99	0.3938
17	REP 1	-3.20886801	3.01031216	3	-1.07	0.3646
18	BLOCK(REP) 2 1	-5.37418373	2.63153067	3	-2.04	0.1338
19	NEW*TREAT 4	-5.84282288	2.78641001	3	-2.10	0.1269

```
options ls = 76;
data augbibd;
   infile 'augbibd.sas';
   input yield rep block treat;
   if (treat > 6) then new = 0; else new = 1;
   if (new) then trtn = 999; else trtn = treat;
proc mixed data = augbibd info nobounds;
   class rep block treat trtn;
   model yield = trtn / solution;
  random rep block(rep) treat*new / solution;
   lsmeans trtn;
   make 'solutionr' out = sr noprint;
run;
proc sort data = sr;
   by descending est;
proc print;
run;
```